GleSYS

Falkenberg 2024-03-22

INCIDENT REPORT REGARDING PARTIAL VMWARE OUTAGE IN FALKENBERG

GleSYS has released the following Reason for Outage (RFO) report.
This document outlines the background and remediation actions in
response to the partial VMware outage in Falkenberg on March 20th, 2024.

BACKGROUND

Redundancy is paramount to the design of our VMware platform at GleSYS. We strive to deliver redundancy across all levels to ensure the platform is stable and resilient for our customers.

We configure our VMware hypervisors to have multiple redundant paths to the storage infrastructure. It is essential to understand that, per our design, these storage paths are configured to traverse two separate network switches. This ensures there is no single point of failure and that we can withstand failures at multiple levels without any customer impact.

Unfortunately, on a small subset of hypervisors in our VMware platform in Falkenberg, a cabling error resulted in a configuration where EVERY path to the storage traversed BOTH network switches. In other words, a small subset of hypervisors were running in a non-redundant configuration. Consequently, when the network team performed a planned reboot of a network switch, all paths to the storage infrastructure went down on those misconfigured hosts.

As a result of this outage, some virtual machines running on the affected hypervisors got their filesystems locked in read-only mode.

SEQUENCE OF EVENTS (CET)

2024-03-20

09:00   The planned maintenance window begins to perform routine upgrades of switches connected to the VMware platform.

09:50   Upgrade of first network switch commences.

10:20   Upgrade of first network switch complete.

10:25   Upgrade of second network switch commences.

10:45   Upgrade of second network switch complete.

13:10   Upgrade of third network switch commences.

13:12   Our monitoring system alerts us that some VMs running in VMware Falkenberg are no longer responding, and there appear to be storage-related issues.

13:30   Upgrade of third network switch complete. Alarms start clearing, and all connectivity appears to be restored.

13:33   We have identified the root cause of the outage. The hosts identified as being misconfigured are running with reduced redundancy.

13:40   Maintenance window is put on hold.

2024-03-21

09:10   We have resolved the underlying misconfiguration. Redundancy is now functioning as intended.

CONCLUSIONS

We will take action to identify and investigate these misconfigurations in a controlled environment. We are also reviewing our planned maintenance and window scheduling procedures and aim to improve our communication on this type of maintenance via https://status.glesys.com/.

Sincerely,

Stefan de Vries
Head of Operations
GleSYS